Comparativo Plataformas de datos abiertos

2015

WORLD BANK GROUP

SOCRATA

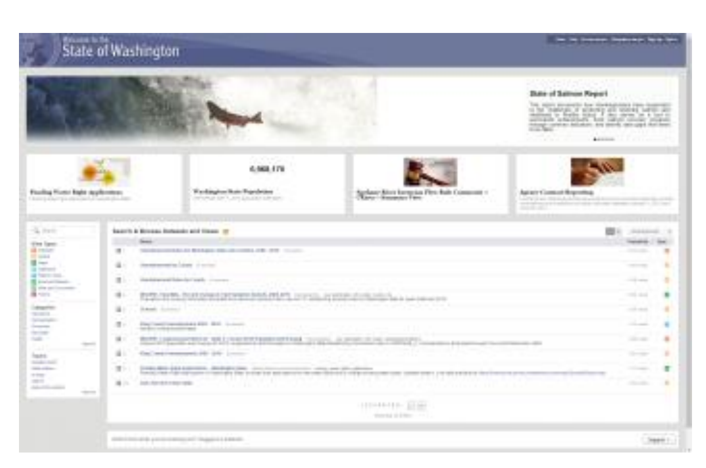
- Socrata es una plataforma de catálogo de datos abiertos tipo SaaS basada en la nube que proporciona herramientas API, catálogo y manipulación de datos. Una característica distintiva de Socrata es que permite a los usuarios crear vistas y visualizaciones basadas en datos publicados y guardarlas para que los demás las puedan utilizar. Además, Socrata ofrece una versión de código abierto de su API, destinada a facilitar la transición para los clientes que decidan migrar del modelo SaaS. http://www.socrata.com



Portales con SOCRATA



Nueva York



Washington



San Francisco



Edmonton



	CKAN	SOCRATA
Platform URL	http://ckan.org/	http://www.socrata.com/products/open-data-portal/
Driven/Owned by	Open Knowledge Foundation	Socrata
Open Source?	Yes	Partial (Socrata core backend data platform is developed in Open Source
Licence	AGPL v3.0	Apache v2.0
Programming Language	Python	Scala
Repository URL	https://github.com/ckan/ckan	https://github.com/socrata-platform
Data Preview/Visualisation	Yes	Yes
Data API?	Yes	Yes
API format (XML, JSON, RDF, etc)	JSON, XML, RDF	JSON, CSV, GeoJSON, OData, RDF-XML
Metadata standard(s)	CKAN metadata schema, plus user-defined key-value pairs. Mostly used are DCAT and Dublin Core.	Allows user-defined metadata schemas; provides compliance with the Common Core Metadata Schema; data.json and DCAT.
Example deployment [1]	http://open-data.europa.eu	https://data.cityofnewyork.us/
Example deployment [2]	http://data.gov.uk/	https://data.cityofchicago.org/
Example deployment [3]	http://datos.gob.mx/	https://data.sfgov.org/
Example deployment [4]	http://www.dati.gov.it/	https://finances.worldbank.org/
Example deployment [5]	http://opendata.admin.ch/	https://data.bathhacked.org/
Example deployment [6]	http://oppnadata.se/	https://opendata.bristol.gov.uk/
Example deployment [7]	http://daten.berlin.de/	https://www.dati.lombardia.it/
Example deployment [8]	https://data.hdx.rwlabs.org/	https://data.lacounty.gov/

Source OpenDataMonitor + ODI:
https://docs.google.com/spreadsheets/d/12aptaQdVrd37QH40nbT2YpTZqd2fWe-V8NQXRsU0Leo/edit#gid=1236063647



Comparativa (i)

	SOCRATA	CKAN
Modelo	Software Comercial	Software Open Source
Soporte	SOCRATA (via partner local)	Soporte comercial (OKF, comunidad Open Source, partner local especializado)
Evolutivos	Dependientes de Fabricante	Depende de: -Comunidad de desarrolladores alrededor de CKAN -La iniciativa de OKF -Otros países en liberación de módulos de sus portales.
Puesta en marcha	Inmediata (con soporte de SOCRATA)	Inmediata (con soporte de OKF)
	Full cloud hosting: Gestión de Datos, APIs, Contenido, Permisos a través de una misma plataforma	Versión Express: SaaS con funcionalidad reducida Versión CKAN Portal: full customizable.
	Producto llave en mano (Configuraciones en función de necesidades: Basic, Extended, Enterprise). Existe la edición 'Community' (Open Source).	Producto customizable. Requiere equipo técnico. No solo con experiencia en Python sino también en Gestión de Datasets y ETLs. Organizaciones como OKF pueden ofrecer este soporte (bajo coste)



Comparativa (ii)

	SOCRATA	CKAN
Nuevos catálogos	Coste bajo (nueva instancia de SOCRATA)	Coste bajo (Versión Express. Nueva instancia)
		Coste alto (Versión Portal. Despliegue en Servidor). Requiere apoyo técnico
Upgrades y Mantenimiento	Upgrades dependientes del Fabricante	Upgrades dependientes del gestor del Portal
	Dependencia del fabricante	Dependencia de la Comunidad (u OKF, u otros portales)
Funcionalidad	Funcionalidad suficiente para las necesidades de OD Colombia, ofrecida de serie por la plataforma	Funcionalidad suficiente para las necesidades de OD Colombia, ofrecida de serie por la plataforma. Esta funcionalidad es extensible. Requiere esfuerzo técnico
	Aplicaciones ad hoc (presupuestos, rendición de cuentas)	129 extensiones disponibles (código disponible en Github). Su adaptación requiere esfuerzo técnico
	Existen herramientas que complementan funcionalidad ETL como FME.	Módulos complementarios basados en el catálogo de extensiones.
	Federación (SOC <-> CKAN)	Federación (SOC <-> CKAN)



Comparativa (iii)

	SOCRATA	CKAN
Integración con front-ends	Si, pero debilmente personalizable	Si. Existen modulos probados de integración con Drupal o Wordpress.
Estandares	La edición 'Community' (OSS) usa vocabularios estándar DCAT (W3C). Resto de versiones metadatos configurables por usuario.	Usa DCAT (W3C). Metadatos configurables por usuario.
Linked Open Data	Datos expuestos como RDF, sin embargo, no permite almacenamiento triple store ni posee un SPARQL endpoint.	No, sólo metadata.
Conocimiento técnico	Menor necesidad de conocimientos técnicos para la gestión y mantenimiento	Mayor necesidad de conocimientos técnicos para la gestión, mantenimiento y extensibilidad de la plataforma.
Clientes	Amplia base de clientes muy concentrados en USA (+80 sobre 300). 1 en LA (Monterrey, MX).	Gran adopción y Amplia base de clientes en todo el mundo (116 sobre 300). 11 en LA.



Conclusiones

Ambas plataformas tienen un catálogo de características para gestión de catalogo de datos suficientes para Open Data en Colombia (APIs para descarga y actualización de datasets, automatización de procesos para exponer datos albergados en sistemas legados, federación de catálogos, integración con front-ends, disponibilidad SaaS, feedback de usuarios, etc.).

SOCRATA es un producto llave en mano de fabricante que permite concentrar esfuerzos en la gestión de datasets.

CKAN es una herramienta en constante proceso evolutivo, soportado por una amplia comunidad de desarrolladores, que requiere un esfuerzo técnico añadido para la gestión y mantenimiento de la plataforma.



Ventajas y otras externalidades

- Socrata es un producto de calidad, claro, sencillo y que ofrece rapidez en la puesta a punto al tratarse de un producto llave en mano y permite focalizar esfuerzos en la gestión de datasets.
- CKAN o DKAN (versión Drupal), conlleva un costo más ventajoso por ser un producto de modelo abierto. Éste requiere mayor atención a los aspectos técnicos (equipo de 2/3 desarrolladores + diseñador), que varían desde albergar internamente las capacidades técnicas para dar respuesta a las necesidades de soporte técnico (desarrollos y adaptaciones in-house), hasta dedicar los esfuerzos a gestionar y dirigir la relación (contractual o no) con la comunidad local de desarrolladores u organizaciones como OKF.
- La pertenencia a la comunidad global de CKAN, permite que los desarrollos de nuevas instancias que se lleven a cabo en otros países u organizaciones como OKF, puedan ser adaptados en otros lugares.
- De cara a la escalabilidad a otras entidades públicas (por ejemplo entidades territoriales), por un lado Socrata sigue ofreciendo una solución sencilla y ágil en la implementación. Por otro lado el fortalecimiento de la comunidad local de desarrolladores aumenta la capacidad de escalabilidad del modelo código abierto a otras entidades, donde el factor de eficiencia de los recursos económicos puede constituir un factor crítico de decisión para sumarse a OD.



Riesgos

SOCRATA	CKAN
Los riesgos asociados a la plataforma SOCRATA son los inherentes al software propietario relacionados a la **dependencia del fabricante**. Sin embargo, la amplia base de instalaciones SOCRATA en funcionamiento lo minimizan. Por otro lado, el hecho de que los evolutivos de la plataforma respondan al único impulso del fabricante pueden llevar a una obsolescencia anticipada de la plataforma por falta de capacidad de adaptación a la evolución de las plataformas tecnológicas Open Data.	Si bien un modelo de código abierto conlleva ventajas en términos de costos, el principal riesgo asociado a CKAN es la imperativa necesidad de **dedicación de recursos técnicos** para gestionar e incorporar funcionalidades a la plataforma y no sólo a la gestión y calidad de los datasets publicados. Esta circunstancia puede encarecer el coste de apropiación del sistema. Por otro lado, este riesgo se puede convertir en una oportunidad combinando los esfuerzos necesarios para **potenciar una comunidad local** que contribuya al proceso de mejora. Es posible que CKAN se adapte mejor a la evolución de las necesidades del mercado de plataformas Open Data.

